

A tie-break model for graph search

Derek G. Corneil ^{*} Jérémie Dusart [†] Michel Habib [†]
Antoine Mamcarz [†] Fabien de Montgolfier [†]

January 27, 2015

Abstract

In this paper, we consider the problem of the recognition of various kinds of orderings produced by graph searches. To this aim, we introduce a new framework, the Tie-Breaking Label Search (TBLS), in order to handle a broad variety of searches. This new model is based on partial orders defined on the label set and it unifies the General Label Search (GLS) formalism of Krueger, Simonet and Berry (2011), and the “pattern-conditions” formalism of Corneil and Krueger (2008). It allows us to derive some general properties including new pattern-conditions (yielding memory-efficient certificates) for many usual searches, including BFS, DFS, LBFS and LDFS. Furthermore, the new model allows easy expression of multi-sweep uses of searches that depend on previous (search) orderings of the graph’s vertex set.

Keywords: Graph search model, tie-break mechanisms, BFS, DFS, LBFS, LDFS, multi-sweep algorithms.

1 Introduction

A graph search is a mechanism for systematically visiting the vertices of a graph. It has been a fundamental technique in the design of graph algorithms since the early days of computer science. Many of the early search methods were based on Breadth First Search (BFS) or Depth First Search (DFS) and resulted in efficient algorithms for practical problems such as the distance between two vertices, diameter, connectivity, network flows and the recognition of planar graphs see [3].

Many variants of these searches have been introduced since, providing elegant and simple solutions to many problems. For example, Lexicographic Breadth First Search (LBFS) [19], and its generalization Maximal Neighbourhood Search (MNS) [21] were shown to yield simple linear time algorithms for chordal graph recognition. More recently,

^{*}Dept. of Computer Science, University of Toronto

[†]LIAFA, UMR 7089 CNRS & Université Paris Diderot, F-75205 Paris Cedex 13, France

Lexicographic Depth First Search (LDFS), was introduced in [8] based on its symmetrical “pattern-condition” with LBFS. A few years after its discovery it was shown that LDFS when applied to cocomparability graphs yields simple and efficient algorithms for solving various Hamiltonian Path related problems [6, 18, 7].

Some recent applications of graph searches involve a controlled tie-break mechanism in a series of consecutive graph searches, see [5, 9, 6, 11]. Examples include the strongly connected components computation using double-DFS [20] and the series of an arbitrary LBFS followed by two LBFS⁺s used to recognize unit interval graphs [5]. Note that a “+ search” breaks ties by choosing (amongst the tied vertices) the vertex that is rightmost with respect to a given ordering of the vertices. This motivates a general study of these graph searches equipped with a tie-break mechanism that incorporates such multi-sweep usage of graph searches. This is the goal of the present paper: to define the simplest framework powerful enough to capture many graph searches either used individually or in a multi-sweep fashion and simple enough to allow general theorems on graph searches. Building on the General Label Search (GLS) framework from [15] we not only simplify their model but also unify their model with the “pattern-conditions” formalism of [8].

This paper is organised as follows. After basic notations and definitions in Section 2, Section 3 introduces the Tie-Breaking Label Search (TBLS) formalism to address graph searches. We then illustrate the TBLS by expressing some classical graph searches in this formalism. We will also show the relationship between our formalism and the “pattern-conditions” of search orderings introduced in [8] thereby yielding some new pattern-conditions for various classical searches. In Section 5 we show that the TBLS and GLS models capture the same set of graph searches. We then propose a unified method for recognizing whether a given ordering of the vertices could have been produced by a specific graph search. Finally, in Section 7 we present a TBLS implementation framework in a particular case.

2 Preliminaries and notation

In this paper, $G = (V, E)$ always (and sometimes implicitly) denotes a graph with n vertices and m edges. All graphs considered here are supposed to be finite. We identify the vertex-set with $\{1, \dots, n\}$, allowing us to see a total ordering on V as a permutation on $\{1, \dots, n\}$.

We define a *graph search* to be an algorithm that visits all the vertices of a graph according to some rules, and a *search ordering* to be the ordering σ of the vertices yielded by such an algorithm. The link between these two notions is an overriding theme of this paper. Vertex $\sigma(i)$ is the i th vertex of σ and $\sigma^{-1}(x) \in \{1, \dots, n\}$ is the position of vertex x in σ . A vertex u is the *leftmost* (respectively *rightmost*) vertex with property X in σ if there is no vertex v such that $X(v)$ and $v <_{\sigma} u$ (respectively $u <_{\sigma} v$). Our graphs are assumed to be undirected, but most searches (especially those captured by TBLS) may be

performed on directed graphs without any modifications to the algorithm (if xy is an arc then we say that y is a neighbour of x while x is *not* a neighbour of y).

The symmetric difference of two sets A and B , namely $(A - B) \cup (B - A)$ is denoted by $A \triangle B$. Furthermore, \mathbb{N}^+ represents the set of integers strictly greater than 0 and \mathbb{N}_p^+ represents the set of integers strictly greater than 0 and less than p . $P(\mathbb{N}^+)$ denotes the power-set of \mathbb{N}^+ and $P_f(\mathbb{N}^+)$ denotes the set of all finite subsets of \mathbb{N}^+ . By \mathfrak{S}_n we denote the set of all permutations of $\{1, \dots, n\}$. For finite $A \in P(\mathbb{N}^+)$, let $umin(A)$ be: if $A = \emptyset$ then $umin(A) = \infty$ else $umin(A) = \min\{i \mid i \in A\}$; and let $umax(A)$ be: if $A = \emptyset$ then $umax(A) = 0$ else $umax(A) = \max\{i \mid i \in A\}$. We always use the notation $<$ for the usual strict (i.e., irreflexive) order between integers, and \prec for a partial strict order between elements from $P_f(\mathbb{N}^+)$ (or from another set when specified). Definitions of most of the searches we will consider appear in [8] or [12].

3 TBLS, a Tie-Breaking Label Search

A *graph search* is an iterative process that chooses at each step a vertex of the graph and numbers it (from 1 to n). Each vertex is chosen (also said *visited*) exactly once (even if the graph is disconnected). Let us now define a general *Tie-Breaking Label Search* (TBLS). It uses *labels* to decide the next vertex to be visited; $label(v)$ is a subset of $\{1, \dots, n\}$. A TBLS is defined on:

1. A graph $G = (V, E)$ on which the search is performed;
2. A strict partial order \prec over the label-set $P_f(\mathbb{N}^+)$;
3. An ordering τ of the vertices of V called the *tie-break permutation*.

The output of $TBLS(G, \prec, \tau)$ is a permutation σ of V , called a *TBLS – ordering* or also the *search ordering* or *visiting ordering*. Let us say a vertex v is *unnumbered* until $\sigma(i) \leftarrow v$ is performed, and then i is its *visiting date*. Thanks to the following algorithm, $label(v)$ is always the set of visiting dates of the neighbours of v visited before v . More specifically $label_i(v)$ for a vertex v denotes the label of v at the beginning of step i . This formalism identifies a search with the orderings it may produce, as in [8], while extending the formalism of General Label Search (GLS) of [15] by the introduction of a *tie-break* ordering τ , making the result of a search algorithm purely deterministic (no arbitrary decision is taken).

Remarks on this formalism:

1. Notice that during a TBLS search vertices are always labelled from 1 up to n . The original description of LBFS generated labels from n down to 1. Since a label is always

Algorithm 1: TBLS(G, \prec, τ)

```
foreach  $v \in V$  do  $label(v) \leftarrow \emptyset$  for  $i \leftarrow 1$  to  $n$  do  
     $Eligible \leftarrow \{x \in V \mid x \text{ unnumbered and } \nexists \text{ unnumbered}$   
     $y \in V \text{ such that } label(x) \prec label(y)\};$   
    Let  $v$  be the leftmost vertex of  $Eligible$  according to the ordering  $\tau$ ;  
     $\sigma(i) \leftarrow v$ ;  
    foreach unnumbered vertex  $w$  adjacent to  $v$  do  
         $label(w) \leftarrow label(w) \cup \{i\};$ 
```

an unordered set rather than a string, as often seen with LBFS and LDFS, we avoid having to prepend or append elements to existing labels. It should also be noticed that the TBLS model does not require the graph to be connected, and therefore in the following we will extend classical graph searches to disconnected graphs. Since we just need to specify \prec to describe a particular search no implementation details have to be discussed in the specification of the search. Finally, by requiring a tie-breaking permutation τ we have a mechanism for choosing a specific vertex from *Eligible*. Many existing recognition algorithms such as the unit interval recognition algorithm in [5] use a series of LBFS sweeps where ties are broken by choosing the rightmost eligible vertex in the previous LBFS search; to accomplish this in the TBLS formalism τ is set to be the reverse of the previous LBFS ordering.

2. Note that all elements of the set *Eligible* have a label set which is maximal with respect to some finite partial order, since any finite partial order has at least one maximal element; therefore *Eligible* cannot be empty. In the context of LBFS the *Eligible* set is often called a *slice*. The reader should be aware that we make no claims on the complexity of computing the strict partial order \prec over the label-set $P_f(\mathbb{N}^+)$; unfortunately it could be NP-hard.

Let us first present an easy characterization property of TBLS search that will be used throughout the paper and which is a direct translation of the algorithm. First we define $N_\sigma(u, v)$ to be the set of visiting dates of neighbours of u that occur before v in σ ; formally $N_\sigma(u, v) = \{i \mid \sigma(i) \in N(u) \text{ and } \sigma(i) <_\sigma v\}$.

Property 3.1. *Let S be a TBLS search with partial order relation \prec_S .*

An ordering σ of the vertices of a graph G is an S -ordering if and only if for every $x, y \in V$, if $x <_\sigma y$, then $N_\sigma(x, x) \not\prec_S N_\sigma(y, x)$.

Proof. The forward direction follows directly from the definition.

For the backward direction, assume that for every $x, y \in V$, $x <_\sigma y$, $N_\sigma(x, x) \not\prec_S N_\sigma(y, x)$ but σ is not an S -ordering. Let $\gamma = \text{TBLS}(G, \prec_S, \sigma)$. Since σ is not an S -ordering, we know that $\sigma \neq \gamma$. Now let i be the first index such that $\sigma^{-1}(i) \neq \gamma^{-1}(i)$.

Let $x = \sigma^{-1}(i)$ and $y = \gamma^{-1}(i)$. Since $x <_\sigma y$ but $\text{TBLS}(G, \prec_S, \sigma)$ did not choose x , we know that x did not have a maximal label at step i . Therefore there must exist z such that $x <_\sigma z$, and $\text{label}_i(x) \prec_S \text{label}_i(z)$. But since $\text{label}_i(x) = N_\sigma(x, x)$ and $\text{label}_i(z) = N_\sigma(z, x)$ we now have a pair of vertices x, z that contradicts the assumption that for all $x, y \in V$, $x <_\sigma y$, $N_\sigma(x, x) \not\prec_S N_\sigma(y, x)$. \square

With this formalism, in order to specify a particular search we just need to specify \prec , the partial order relation on the label sets for that search. As a consequence we can transmit relationships between partial orders to their associated graph searches.

There are two natural ways of saying that a search S is a search S' (for instance, that LBFS is a BFS): either the \prec ordering used by S is a refinement of that of S' ; or any search ordering σ output by an execution of S could also have been output by an execution of S' . In fact it can easily be shown that both formulations are equivalent, as stated in Theorem 3.5.

Definition 3.2. For two TBLS searches S, S' , we say that S' is an extension of S (denoted by $S \ll S'$) if and only if every S' -ordering σ also is an S -ordering.

The statement and proof of the next lemma follows the work of [15] where there are similar results for the GLS formalism.

Lemma 3.3 (see [15]). For any TBLS S , any integer $p \geq 1$ and any sets A and B of $P(\mathbb{N}_p^+)$, if $A \not\prec_S B$ then there exists a graph G and an S -ordering σ such that in the $(p-1)$ st step the label of the $(p-1)$ st vertex is A and the label of the p th vertex is B (i.e., $\text{label}_{p-1}(\sigma^{-1}(p-1)) = A$ and $\text{label}_{p-1}(\sigma^{-1}(p)) = B$).

Proof. Let $G = (V, E)$, with $V = \{z_1, \dots, z_p\}$ and $E = \{z_i z_k \mid 1 \leq k < i \leq p-2 \text{ and if } A \cap \mathbb{N}_i^+ \prec_S B \cap \mathbb{N}_i^+ \text{ then } k \in B \text{ else } k \in A\} \cup \{z_{p-1} z_k \mid k \in A\} \cup \{z_p z_k \mid k \in B\}$. Let $\sigma = (z_1, \dots, z_p)$. By the definitions of E and σ , for any integers i, j such that $1 \leq i \leq j \leq p$, $N_\sigma(\sigma^{-1}(j), \sigma^{-1}(i)) = A \cap \mathbb{N}_i^+$ or $N_\sigma(\sigma^{-1}(j), \sigma^{-1}(i)) = B \cap \mathbb{N}_i^+$ with $N_\sigma(\sigma^{-1}(i), \sigma^{-1}(i)) \not\prec_S N_\sigma(\sigma^{-1}(j), \sigma^{-1}(i))$. By Property 3.1, σ is an S ordering and we have $N_\sigma(\sigma^{-1}(p-1), \sigma^{-1}(p-1)) = A$ and $N_\sigma(\sigma^{-1}(p), \sigma^{-1}(p-1)) = B$. \square

Definition 3.4. For two partial orders \prec_P, \prec_Q on the same ground set X , we say that \prec_P is an extension of \prec_Q if $\forall x, y \in X$, $x \prec_Q y$ implies $x \prec_P y$.

Theorem 3.5. Let S, S' be two TBLS. S' is an extension of S if and only if $\prec_{S'}$ is an extension of \prec_S .

Proof. For the forward direction, assume for contradiction that S' is an extension of S but $\prec_{S'}$ is not an extension of \prec_S . Therefore there exists A, B such that $A \prec_S B$ and $A \not\prec_{S'} B$. Now using Lemma 3.3 there exists a graph G and an S' -ordering σ such that

$label_{p-1}(\sigma^{-1}(p-1)) = A$ and $label_{p-1}(\sigma^{-1}(p)) = B$. Since $A \prec_S B$ using Property 3.1, we deduce that σ is not an S -ordering which contradicts the fact that S' is an extension of S .

Suppose that σ is an S' -ordering. Therefore using Property 3.1 we know that for every $x, y \in V$, $x <_\sigma y$, we have $N_\sigma(x, x) \not\prec_{S'} N_\sigma(y, x)$. Since $\prec_{S'}$ is an extension of \prec_S , we deduce that $x, y \in V$, such that for every $x <_\sigma y$, we have $N_\sigma(x, x) \not\prec_S N_\sigma(y, x)$. Now using Property 3.1, we deduce that σ is an S -ordering. □

The choice of permutation τ is useful in some situations described below; otherwise, we consider the orderings output by an arbitrary choice of τ thanks to the following definition:

Definition 3.6. *Let \prec be some ordering over $P_f(\mathbb{N}^+)$. Then σ is a TBLS ordering for G and \prec if there exists τ such that $\sigma = TBLS(G, \prec, \tau)$.*

Before giving some examples of appropriate \prec for well known searches, let us start with a kind of fixed point theorem and some general features of the TBLS formalism.

Theorem 3.7. *Let G be a graph; \prec a search rule; and σ an ordering of the vertices of G . Then there exists τ such that $\sigma = TBLS(G, \prec, \tau)$ if and only if $\sigma = TBLS(G, \prec, \sigma)$.*

Proof. One direction is obvious. For the other direction, assume that $\sigma = TBLS(G, \prec, \tau)$ for some τ , and consider $\sigma' = TBLS(G, \prec, \sigma)$. Assume, by contradiction, that $\sigma \neq \sigma'$, and consider i , the index of the first difference between σ and σ' . Let $Eligible_i^\sigma$ be the set of eligible vertices at step i of the algorithm that generated σ , and let $Eligible_i^{\sigma'}$ be the set of eligible vertices at step i of the algorithm that generated σ' . Since σ and σ' are equal until index i , $Eligible_i^\sigma = Eligible_i^{\sigma'}$. By the definition of TBLS, $\sigma(i)$ is the first vertex of $Eligible_i^\sigma$ according to τ . Finally, since the first vertex of this set, according to σ , is $\sigma(i)$, the tie-break rule chose it and so $\sigma(i) = \sigma'(i)$, a contradiction. □

This easy result (which is a direct consequence of our model equipped with a built-in tie-break process) is in fact a very powerful theoretical tool to show that some ordering is not a TBLS ordering, and we will use it several times in the proofs in the next sections, as for example in Theorem 4.2.

As a first conclusion, the TBLS model is a pure mathematical abstraction of graph search algorithms via partial orders but with no data structures involved. Moreover, if we have a characterization of the total orderings produced by a given TBLS (as for example the usual search characterizations of section 4) then we can get rid of the implementation itself which can be parallel or sequential. In the next sections we will exhibit some easy consequences of this model. But before, we demonstrate its expressive power, in particular to deal with multi-sweep algorithms (i.e., algorithms written as a series of successive graph searches).

To this aim, let us consider the sequence $\{\sigma_i\}_{i \in \mathbb{N}}$ of total orderings of the vertices, satisfying the following recursive equations:

- (i) σ_0 is an arbitrary total ordering of the vertices
- (ii) $\sigma_i = \text{TBLS}(G, \prec, \sigma_{i-1}^r)$ where σ^r denotes the reverse ordering of σ .

It was proven when \prec is the partial order associated to the LBFS search, as described in the next section:

- (i) in [5], that if G is a unit interval graph then σ_3 is a unit interval ordering¹.
- (ii) in [11] that if G is a cocomparability graph then $\sigma_{|V|}$ is a cocomp ordering².

4 Characterizing classical searches using TBLS

In this section we show how various classical searches (see [8] for the definitions of these various searches) may be expressed in the TBLS formalism. In each case we will state an appropriate \prec order and where applicable, we will establish various characterizations of the search including the “pattern-condition” presented in [8]. In many cases we will exhibit new vertex ordering characterizations.

Definition 4.1. *For every vertex x , let $ln(x)$ be the leftmost (in σ) left neighbour of x , and let $rn(x)$ be the rightmost (in σ) right neighbour of x . In both cases, if x has no left (respectively right) neighbour, then $ln(x)$ (respectively $rn(x)$) = -1 .*

4.1 Generic Search

A *Generic Search* as described by Tarjan [22] is any search that wherever possible visits neighbours of already visited vertices (this corresponds to the usual notion of graph search).

We now give an alternative proof based on our formalism of the characterization of a generic search ordering, called a GEN-ordering throughout the rest of the paper.

Theorem 4.2 (see [8]). *We define $A \prec_{gen} B$ if and only if $A = \emptyset$ and $B \neq \emptyset$ and let σ be a permutation of V . The following conditions are equivalent:*

1. *Vertex ordering σ is a GEN-ordering of V (i.e., a TBLS using \prec_{gen}).*
2. *For every triple of vertices a, b, c such that $a <_\sigma b <_\sigma c$, and $a \in N(c) - N(b)$ there exists $d \in N(b)$ such that $d <_\sigma b$.*
3. *For every $x \in V$, for every $y \in V$ such that $x <_\sigma y <_\sigma rn(x)$, we have $ln(y) \neq -1$.*

Proof. Suppose that σ is a GEN-ordering. Using Property 3.1 on σ , we know that: σ is a GEN-ordering

\iff for every $x, y \in V$ such that $x <_\sigma y$, we have $N_\sigma(x, x) \not\prec_{gen} N_\sigma(y, x)$

¹an ordering τ of V such that for all $x <_\tau y <_\tau z$ with $xz \in E$, we have $xy, yz \in E$.

²an ordering τ of V such that for all $x <_\tau y <_\tau z$ with $xz \in E$, we have at least one of $xy, yz \in E$.

\iff for every $x, y \in V$ such that $x <_\sigma y$, we have $N_\sigma(y, x) = \emptyset$ or $N_\sigma(x, x) \neq \emptyset$
 \iff for every $x, y \in V$ such that $x <_\sigma y$, we have $N_\sigma(y, x) \neq \emptyset \Rightarrow N_\sigma(x, x) \neq \emptyset$
 \iff for every triple of vertices a, b, c such that $a <_\sigma b <_\sigma c$, $a \in N(c) - N(b)$, there exists $d \in N(b)$ such that $d <_\sigma b$.

Therefore we proved the equivalence between 1 and 2. Let us consider 3, which is a reformulation of 2. The fact that $1 \Rightarrow 3$, is obvious. To prove the converse, we can use Theorem 3.7. Suppose that there exists σ satisfying 3 but not 1. Let $\sigma' = \text{TBLs}(G, \prec_{\text{gen}}, \sigma) \neq \sigma$ and i be the leftmost index where they differ ($z = \sigma'(i) \neq y = \sigma(i)$). This means that $l_i(y) = \emptyset$ and there exists some $x \in l_i(z)$. But this implies with condition 3 since $x <_\sigma y <_\sigma z \leq_\sigma rn(x)$ that $ln(y) \neq -1$ contradicting $l_i(y) = \emptyset$. \square

Remark 4.3. *Theorem 3.7 can be used also in the following proofs of this section, but we omit them to avoid tedious reading.*

4.2 BFS (Breadth-First Search)

We now focus on BFS. Here, we will follow the definition of BFS given in [8], that is a graph search in which the vertices that are eligible are managed with a queue. Note that this differs for example from the definition given in [3], where BFS stands for what we call layered search. Our notion of BFS is the most common implementation of a layered search.

Theorem 4.4. *We define $A \prec_{\text{BFS}} B$ if and only if $umin(A) > umin(B)$. Let σ be a permutation of V . The following conditions are equivalent:*

1. *Vertex ordering σ is a BFS-ordering (i.e., a TBLs using \prec_{BFS}).*
2. *For every triple $a, b, c \in V$ such that $a <_\sigma b <_\sigma c$, $a \in N(c) - N(b)$, there exists d such that $d \in N(b)$ and $d <_\sigma a$.*
3. *For every triple $a, b, c \in V$ such that $a <_\sigma b <_\sigma c$ and a is the leftmost vertex of $N(b) \cup N(c)$ in σ , we have $a \in N(b)$.*

Proof. The equivalence between condition (1) and condition (2) has been proved in [8]. We now prove that condition (1) is equivalent to condition (3). Suppose that σ is a BFS-ordering. Using Property 3.1 on σ , we know that:

σ is a BFS-ordering

\iff for every $x, y \in V$, $x <_\sigma y$, $N_\sigma(x, x) \not\prec_{\text{BFS}} N_\sigma(y, x)$
 \iff for every $x, y \in V$, $x <_\sigma y$, $umin(N_\sigma(x, x)) \not\prec umin(N_\sigma(y, x))$
 \iff for every $x, y \in V$, $x <_\sigma y$, $umin(N_\sigma(x, x)) \leq umin(N_\sigma(y, x))$
 \iff for every triple of vertices a, b, c such that $a <_\sigma b <_\sigma c$, and a is the leftmost vertex of $N(b) \cup N(c)$, we have $a \in N(b)$.

\square

4.3 DFS (Depth First Search)

We now turn our attention to Depth First Search.

Theorem 4.5. *We define $A \prec_{DFS} B$ if and only if $umax(A) < umax(B)$. Let σ be a permutation of V . The following conditions are equivalent:*

1. *Vertex ordering σ is a DFS-ordering (i.e., a TBLS using \prec_{DFS}).*
2. *For every triple of vertices a, b, c such that $a <_{\sigma} b <_{\sigma} c$, $a \in N(c) - N(b)$ there exists $d \in N(b)$ such that $a <_{\sigma} d <_{\sigma} b$.*
3. *For every triple of vertices a, b, c such that $a <_{\sigma} b <_{\sigma} c$, and a is the rightmost vertex of $N(b) \cup N(c)$ to the left of b in σ , we have $a \in N(b)$.*

Proof. The equivalence between condition (1) and (2) has been proved in [8]. Let us show the equivalence between (1) and (3). Suppose that σ is a DFS-ordering. Using Property 3.1 on σ , we know that:

σ is a DFS-ordering

- \iff for every $x, y \in V$ such that $x <_{\sigma} y$, we have $N_{\sigma}(x, x) \not\prec_{DFS} N_{\sigma}(y, x)$
- \iff for every $x, y \in V$ such that $x <_{\sigma} y$, we have $umax(N_{\sigma}(x, x)) \not\leq umax(N_{\sigma}(y, x))$
- \iff for every $x, y \in V$ such that $x <_{\sigma} y$, we have $umax(N_{\sigma}(y, x)) \leq umax(N_{\sigma}(x, x))$
- \iff for every triple of vertices a, b, c such that $a <_{\sigma} b <_{\sigma} c$ and a is the rightmost vertex of $N(b) \cup N(c)$ to the left of b in σ , we have $a \in N(b)$.

□

4.4 LBFS (Lexicographic Breadth First Search)

LBFS was first introduced in [19] to recognize chordal graphs. Since then many new applications of LBFS have been presented ranging from recognizing various families of graphs to finding vertices with high eccentricity or to finding the modular decomposition of a given graph, see [13, 4, 9, 24].

Theorem 4.6. *We define $A \prec_{LBFS} B$ if and only if $umin(B - A) < umin(A - B)$. Let σ be a permutation of V . The following conditions are equivalent:*

1. *Vertex ordering σ is a LBFS-ordering (i.e., a TBLS using \prec_{LBFS}).*
2. *For every triple $a, b, c \in V$ such that $a <_{\sigma} b <_{\sigma} c$, $a \in N(c) - N(b)$, there exists $d <_{\sigma} a$, $d \in N(b) - N(c)$.*
3. *For every triple $a, b, c \in V$ such that $a <_{\sigma} b <_{\sigma} c$ and a is the leftmost vertex of $N(b) \triangle N(c)$ to the left of b in σ , then $a \in N(b) - N(c)$.*

Proof. The equivalence between (1) and (2) is well known, see [19, 12, 2]. We now prove the equivalence between (1) and (3). Suppose that σ is a LBFS-ordering. Using Property 3.1 on σ , we know that:

σ is a LBFS-ordering

- \iff for every $x, y \in V$, $x <_\sigma y$, we have $N_\sigma(x, x) \not\prec_{LBFS} N_\sigma(y, x)$
- \iff for every $x, y \in V$, $x <_\sigma y$, we have $umin(N_\sigma(y, x) - N_\sigma(x, x)) \not\prec umin(N_\sigma(x, x) - N_\sigma(y, x))$
- \iff for every $x, y \in V$, $x <_\sigma y$, we have $umin(N_\sigma(x, x) - N_\sigma(y, x)) \leq umin(N_\sigma(y, x) - N_\sigma(x, x))$
- \iff for every triple of vertices a, b, c such that $a <_\sigma b <_\sigma c$ and a is the leftmost vertex of $N(b) \triangle N(c)$ to the left of b in σ , we have $a \in N(b) - N(c)$.

□

4.5 LDFS (Lexicographic Depth First Search)

Lexicographic Depth First Search (LDFS) was introduced in [8].

Theorem 4.7. *We define $A \prec_{LDFS} B$ if and only if $umax(A - B) < umax(B - A)$. Let σ be a permutation of V . The following conditions are equivalent:*

1. *Vertex ordering σ is a LDFS-ordering (i.e., a TBLS using \prec_{LDFS})*
2. *For every triple $a, b, c \in V$ such that $a <_\sigma b <_\sigma c$, $a \in N(c) - N(b)$, there exists $a <_\sigma d <_\sigma b$, $d \in N(b) - N(c)$.*
3. *For every triple $a, b, c \in V$ such that $a <_\sigma b <_\sigma c$ and a is the rightmost vertex in $N(b) \triangle N(c)$ to the left of b in σ , $a \in N(b) - N(c)$.*

Proof. The equivalence between (1) and (2) is well known, see [8]. We now prove the equivalence between (1) and (3). Suppose that σ is a LDFS-ordering. Using Property 3.1 on σ , we know that:

σ is a LDFS-ordering

- \iff for every $x, y \in V$ such that $x <_\sigma y$, we have $N_\sigma(x, x) \not\prec_{LDFS} N_\sigma(y, x)$
- \iff for every $x, y \in V$ such that $x <_\sigma y$, we have $umax(N_\sigma(x, x) - N_\sigma(y, x)) \not\prec umax(N_\sigma(y, x) - N_\sigma(x, x))$
- \iff for every $x, y \in V$ such that $x <_\sigma y$, we have $umax(N_\sigma(y, x) - N_\sigma(x, x)) \leq umax(N_\sigma(x, x) - N_\sigma(y, x))$
- \iff for every triple of vertices a, b, c such that $a <_\sigma b <_\sigma c$ and a is the rightmost vertex of $N(b) \triangle N(c)$ to the left of b in σ , we have $a \in N(b) - N(c)$.

□

The symmetry between BFS and DFS (respectively LBFS and LDFS) becomes clear when using the TBLS ordering formalism. This symmetry was also clear using the pattern-conditions as introduced in [8], and in fact lead to the discovery of LDFS.

To finish with classical searches, we notice that **Maximum Cardinality Search (MCS)** as introduced in [23], can easily be defined using the partial order: $A \prec_{MCS} B$ if and only if $|A| < |B|$. Similarly **MNS (Maximal Neighbourhood Search)** as introduced in [21] for chordal graph recognition, is a search such that $A \prec_{MNS} B$ if and only if $A \subsetneq B$, i.e., it uses the strict inclusion partial order between subsets.

To conclude this section we use Theorem 3.5 to easily rediscover the relationships amongst various graph searches as noted in [8] and [15].

Theorem 4.8. *The partial order of the relation extension between classical searches is described in Figure 1.*

Proof. To show that a search extends another one we will use Theorem 3.5.

Let us show that \prec_{BFS} (respectively \prec_{DFS} , \prec_{MNS}) is an extension of \prec_{gen} . Let $A \prec_{gen} B$. By definition we have $A = \emptyset$ and $B \neq \emptyset$. As a consequence we have $umin(B) < umin(A)$ and thus $A \prec_{BFS} B$ (respectively $umax(A) < umax(B)$ implying $A \prec_{DFS} B$, and $A \subsetneq B$ implying $A \prec_{MNS} B$).

We now show that \prec_{LBFS} is an extension of \prec_{BFS} . Let $A \prec_{BFS} B$. By definition, we have $umin(B) < umin(A)$. As a consequence, $umin(B - A) < umin(A - B)$, implying $A \prec_{LBFS} B$.

To see that \prec_{LDFS} is an extension of \prec_{DFS} , first suppose that $A \prec_{DFS} B$. Therefore $umax(A) < umax(B)$ and as a consequence $umax(A - B) < umax(B - A)$ thereby implying $A \prec_{LDFS} B$.

Similarly \prec_{LBFS} (respectively \prec_{LDFS} , \prec_{MCS}) is an extension of \prec_{MNS} . Let $A \prec_{MNS} B$; by definition we have $A \subsetneq B$. As a consequence, $umin(B - A) < umin(A - B)$ (respectively $umax(A - B) < umax(B - A)$ and $|A| < |B|$). So $A \prec_{LBFS} B$ (respectively $A \prec_{LDFS} B$ and $A \prec_{MCS} B$).

□

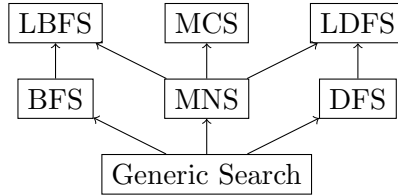


Figure 1: Summary of the hereditary relationships proved in Theorem 4.8. An arc from Search S to Search S' means that S' extends S .

4.6 Limitations of the TBLS model

To finish, let us remark that there exists at least one known search that does not fit into the TBLS model. In the following, recall that $label_i(v)$ for a vertex v denotes the label of v at the beginning of step i of Algorithm 1. *Layered Search* starts at a vertex s , and ensures that if $dist(s, x) < dist(s, y)$ then $\sigma(x) < \sigma(y)$. In other words it respects the layers (vertices at the same distance from the start vertex s). We now show that this search is not a TBLS by considering the graph G in Figure 2. Assume that we have started the Layered Search with x_1, x_2, x_3, x_4 and so $label_5(x_5) = \{3\}$ and $label_5(x_6) = \{4\}$. In a Layered Search, both x_5 and x_6 must be eligible at step 5. Thus we must have neither $\{3\} \prec \{4\}$ nor $\{4\} \prec \{3\}$; they are incomparable labels. But now consider graph H in Figure 2 and assume that again we have started the search with v_1, v_2, v_3, v_4 . So we have $label_5(v_5) = \{3\}$ and $label_5(v_6) = \{4\}$. But in this graph we have to visit v_5 before v_6 . Therefore we must have $\{3\} \prec \{4\}$. As a conclusion, no partial ordering of the labels can capture all Layered Search orderings and so this search cannot be written in our formalism. The same seems true for Min-LexBFS as defined in [17] and Right Most Neighbour as used in [6].

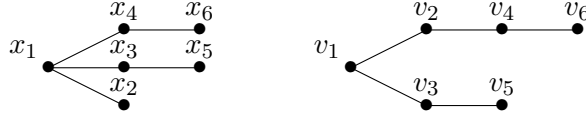


Figure 2: Graph G on the left and H on the right.

5 The relationship between GLS and TBLS

We are now interested in determining the relationship between TBLS and GLS. First let us recall GLS from [15]. It depends on a *labeling structure* which consists of four elements:

- a set of labels L ;
- a strict order \prec_{GLS} over the label-set L ;
- an initial label l_0 ;
- an UPLAB function $L \times \mathbb{N}^+ \rightarrow L$.

The GLS algorithm then takes as input a graph $G = (V, E)$ (over which the search is performed) as well as a labeling structure.

The computational power of the UPLAB function is unbounded, even though it must be deterministic, and the label set L may be any set. In contrast, TBLS uses a fixed initial

label \emptyset , a fixed label set $P_f(\mathbb{N}^+)$, and a fixed simple updating function. Despite these restrictions, it is, however equivalent to GLS in the sense of Theorem 5.3.

Algorithm 2: $GLS(G, \{L, \prec_{GLS}, l_0, UPLAB\})$

```

foreach  $v \in V$  do  $l(v) \leftarrow l_0$  for  $i \leftarrow 1$  to  $n$  do
    Let Eligible be the set of eligible vertices, i.e., those unnumbered vertices  $v$  with  $l(v)$ 
    maximal with respect to  $\prec_{GLS}$ ;
    Let  $v$  be some vertex from Eligible;
     $\sigma(i) \leftarrow v$ ;
    foreach unnumbered vertex  $w$  adjacent to  $v$  do
         $l(w) \leftarrow UPLAB(l(w), i)$ ;

```

We now prove that for each GLS , there is a \prec_{TBLS} producing the same orderings, and conversely. First we need some notation.

At each iteration i of $TBLS(G, \prec_{TBLS}, \sigma)$, let $l_{TBLS,i}(v)$ be the label assigned to every unnumbered vertex v by $TBLS(G, \prec_{TBLS}, \sigma)$, i.e., the label that will be used to choose the i th vertex. Similarly, let $l_{GLS,i}(v)$ be the label assigned to every unnumbered vertex v by $GLS(G, \{L, \prec_{GLS}, l_0, UPLAB\})$, i.e., the label that will be used to choose the i th vertex. Given a graph $G = (V, E)$, and an ordering σ of V , let us define $I_k^\sigma(v) = N(v) \cap \{\sigma(1), \dots, \sigma(k)\}$ to be the neighbours of v visited at step k or before. Let us define p_k^j to be the j -th element of $I_k^\sigma(v)$ sorted in increasing visiting ordering.

Proposition 5.1 ([15]). *Let S be a labeling structure, $G = (V, E)$ a graph. At iteration i of $GLS(G, S)$ computing an ordering σ , for every unnumbered vertex v :*
 $l_{GLS,i}(v) = UPLAB(\dots UPLAB(l_0, p_1), \dots, p_k)$ *where (p_1, \dots, p_k) is the sequences of numbers in $N_\sigma(v, \sigma^{-1}(i))$ in increasing order.*

Proposition 5.2. *Let $G = (V, E)$ be a graph, $v \in V$, and σ the ordering produced by $TBLS(G, \prec, \tau)$. At iteration i of $TBLS(G, \prec, \tau)$, for every unnumbered vertex v :*
 $l_{TBLS,i}(v) = I_{i-1}^\sigma(v)$.

Proof. The proof goes by induction. At the first step of the algorithm, every vertex has \emptyset as its label, and has no previously visited neighbour.

Assume that at iteration i , every unnumbered vertex x has label $l_{TBLS,i}(x) = I_{i-1}^\sigma(x)$. After this iteration, for every unnumbered neighbour v of $\sigma(i)$,
 $l_{TBLS,i}(v) = l_{TBLS,i-1}(v) \cup \{i\}$, which is indeed $I_i^\sigma(v)$, and for every unnumbered non-neighbour v of $\sigma(i)$, $l_{TBLS,i}(v) = l_{TBLS,i-1}(v)$, which is again $I_i^\sigma(v)$. □

Theorem 5.3. *A set T of orderings of the vertices of a graph G is equal to $\{TBLS(G, \prec_{TBLS}, \tau) \mid \tau \in \mathfrak{S}_n\}$ if and only if there exists a labeling structure $S = (L, \prec_{GLS}, l_0, UPLAB)$ such that T is equal to the set of orderings produced by $GLS(G, S)$.*

Proof. First, consider an ordering \prec_{TBLS} . The set $\{TBLS(G, \prec_{TBLS}, \tau) \mid \tau \in \mathfrak{S}_n\}$ is equal to the set of all orderings produced by $GLS(G, S)$ with $S = (P(\mathbb{N}^+), \prec_{TBLS}, \emptyset, cons)$, where $cons(l(w), i)$ returns $l(w) \cup \{i\}$.

Conversely, consider $S = (L, \prec_{GLS}, l_0, UPLAB)$ a labeling structure. We show that there exists an order \prec_{TBLS} such that, for every graph G , the set of all orderings produced by $GLS(G, S)$ is equal to $\{TBLS(G, \prec_{TBLS}, \tau) \mid \tau \in \mathfrak{S}_n\}$.

By propositions 5.1 and 5.2 we can define a mapping ϕ from $P(\mathbb{N}^+)$ (the labels used by TBLS) into labels *effectively* used by GLS (i.e., those that can be assigned to a vertex during some execution of the algorithm). ϕ is recursively defined as $\phi(\emptyset) = l_0$, and if $max(A) = i$, then $\phi(A) = UPLAB(\phi(A) \setminus \{i\}, i)$. Notice the same GLS-label l may be reached in different ways. Subset $\phi^{-1}(l) \subset P(\mathbb{N}^+)$ is the set of TBLS-labels that correspond to that label. It is empty for all labels not effectively used.

Then, we define \prec_{TBLS} as follows: $\forall A, A' \in P_f(\mathbb{N}^+)$, $A \prec_{TBLS} A'$ if and only if $\phi(A) \prec_{GBLS} \phi(A')$. We are now ready to prove the theorem. The proof goes by induction. Before the first iteration, for every vertex v , $l_{GLS}^0(v) = l_0$ and $l_{TBLS}^0(v) = \emptyset$. GLS can pick any of these vertices, in particular the one that would be picked by $TBLS(G, \prec_{TBLS}, \tau)$, and by setting τ to be equal to a given output of $GLS(G, S)$, TBLS would indeed chose the same vertex.

Now, assume that when step i begins, both algorithms have produced the ordering $\sigma(1) \dots \sigma(i-1)$, and the i th vertex is about to be chosen. By propositions 5.1 and 5.2, for every unnumbered vertex x , $l_{TBLS,i}(x) = I_{i-1}^\sigma(x)$, and $l_{GLS,i}(x) = UPLAB(\dots UPLAB(l_0, p_{i-1}^1) \dots, p_{i-1}^{|I_k^\sigma(x)|})$. By the definition of ϕ , we have that $l_{GLS,i}(x) = \phi(l_{TBLS,i}(x))$, and $l_{TBLS,i}(x) \in \phi^{-1}(l_{GLS,i}(x))$. Then, by the definition of \prec_{TBLS} , for two unnumbered vertices v and w , we know that $l_{TBLS,i}(v) \prec_{TBLS} l_{TBLS,i}(w)$ if and only if $\phi(l_{TBLS,i}(v)) \prec_{GLS} \phi(l_{TBLS,i}(w))$, and $l_{GLS,i}(v) \prec_{GLS} l_{GLS,i}(w)$ if and only if for all $l_v \in \phi^{-1}(l_{GLS,i}(v))$ and all $l_w \in \phi^{-1}(l_{GLS,i}(w))$, $l_v \prec_{TBLS} l_w$. Thus, the set of eligible vertices at step i is the same for both algorithms. GLS can pick any of these vertices, in particular the one that would be picked by $TBLS(G, \prec_{TBLS}, \tau)$, and by setting τ to be equal to $GLS(G, S)$, TBLS would indeed choose the right vertex. \square

Although TBLS and GLS cover the same set of vertex orderings, we think that our TBLS formalism provides a simpler framework to analyze graph search algorithms, as can be seen in the next section.

6 Recognition of some TBLS search orderings

Let us now consider the following problem:

Recognition of Search \mathcal{S}

Data: Given a total ordering σ of the vertices of a graph G and a TBLS search \mathcal{S} ,

Result: Does there exist τ such that $\sigma = TBLS(G, \prec_{\mathcal{S}}, \tau)$?

Of course we can use Theorem 3.7 and build an algorithm that tests whether or not $\sigma = \text{TBLs}(G, \prec_S, \sigma)$. Let $\tau = \text{TBLs}(G, \prec_S, \sigma)$. If $\tau = \sigma$ then the answer is yes; otherwise it is no. We can certify the no answer using the first difference between τ and σ . Let i be the first index such that $\sigma(i) \neq \tau(i)$. If TBLs chooses $\tau(i)$ and not $\sigma(i)$ at step i , then at this time $l(\sigma(i)) \prec_S l(\tau(i))$. So we can build a contradiction to the pattern-condition of this search.

But we may want to be able to answer this question without applying a TBLs search, or modifying a TBLs algorithm. For example suppose that a distributed or parallel algorithm has been used to compute the ordering (for example when dealing with a huge graph [1]) that is assumed to be a specific search ordering; how does one efficiently answer this question? Let us study some cases.

6.1 Generic Search

For Generic Search consider Algorithm 3 where σ is the ordering we want to check, and for all i between 1 and n , $ln(\sigma(i))$ has been computed; note that G may be disconnected. Recall that $ln(x)$ is the leftmost left neighbour of x ; if x has no left neighbours, then $ln(x) = -1$. The algorithm will output either “YES” or “NO” depending on whether or not σ is a GEN-ordering.

Algorithm 3: GEN-check

```

 $J \leftarrow 1$ ;    % { If  $\sigma$  is a GEN-ordering, then  $J$  is the index of the first
                vertex of the current connected component.}%
for  $i \leftarrow 2$  to  $n$  do
    if  $ln(\sigma(i)) = -1$  then
         $J \leftarrow i$ 
    else
        if  $ln(\sigma(i)) < J$  then
            return “NO”
return “YES”

```

Theorem 6.1. *The GEN-check algorithm is correct and requires $O(n)$ time. The recognition of a GEN-ordering can be implemented to run in $O(n + m)$ time.*

Proof. If the algorithm reports that σ is not a GEN-ordering, then vertices $\sigma(ln(i)), \sigma(J), \sigma(i)$ form a forbidden triple as stipulated in Condition 2 of Theorem 4.2. Note that $\sigma(J)$ has no neighbours to its left in σ .

Now assume that the algorithm reports that σ is a GEN-ordering but for sake of contradiction there exists a forbidden triple on vertices $a <_\sigma b <_\sigma c$. Let J be the rightmost J index less than $\sigma^{-1}(c)$ identified by the algorithm; note that $b \leq_\sigma \sigma(J) <_\sigma c$ and

$\sigma(ln(c)) \leq_\sigma a$. When $i = \sigma^{-1}(c)$ the algorithm would have reported that σ is not a GEN-ordering.

For the preprocessing we need to compute the values of $ln(x)$ for every vertex x , following Definition 4.1. By sorting the adjacency lists with respect to σ (in linear time), it is possible to find $ln(x)$ in linear time by scanning the adjacency lists once and storing $ln(x)$ in an array. Given this information, Algorithm 3 runs in $O(n)$ time. Including the preprocessing time, the whole complexity needed is $O(n + m)$. \square

6.2 BFS

In order to handle the recognition of BFS-orderings and DFS-orderings, we will first prove variations of the conditions proposed in Theorems 4.4 and 4.5, which are easier to check. Let us define for every vertex x in V , the following two intervals in σ : $Right(x) = [x, rn(x)]$ and $Left(x) = [ln(x), x]$. By convention, if $rn(x) = -1$ or $ln(x) = -1$ the corresponding interval is reduced to $[x]$.

Theorem 6.2. *Vertex ordering σ is a BFS-ordering of V if and only if*

1. *Vertex ordering σ is a GEN-ordering of G*
2. *For every pair of vertices x, y , if $x <_\sigma y$ then $ln(x) \leq_\sigma ln(y)$*
3. *For every pair of vertices x, y , if $x \neq y$ then the intervals $Left(x)$ and $Left(y)$ cannot be strictly included*³.

Proof. It is easy to show that Conditions 2 and 3 are equivalent.

\Rightarrow First, notice that every BFS-ordering σ is also a GEN-ordering. Now assume for contradiction that Condition 3 is contradicted, namely that $x <_\sigma y$ and that $Left(y)$ strictly contains $Left(x)$. Then we have the configuration: $ln(y) <_\sigma ln(x) \leq_\sigma x <_\sigma y$. Considering the triple $(ln(y), x, y)$, since $ln(y) <_\sigma ln(x)$, necessarily $xln(y) \notin E$. Using the BFS 4-points condition on this triple there exists z such that $z <_\sigma ln(y)$ where $xz \in E$, thereby contradicting $ln(y) <_\sigma ln(x)$.

\Leftarrow Assume that σ respects all three conditions of the theorem. Consider a triple (a, b, c) of vertices such that: $a <_\sigma b <_\sigma c$ with $ac \in E$ and $ab \notin E$. Since σ is a GEN-ordering, $ac \in E$ implies that $ln(b) \neq -1$ (i.e., b has a left neighbour in σ).

Suppose $ln(b) >_\sigma a$. Since $ln(c) \leq_\sigma a$, this implies that $Left(c)$ strictly contains $Left(b)$, thereby contradicting Condition 3. Therefore b has a neighbour before a in σ . So σ follows the BFS 4-points condition and is a legitimate BFS-ordering. \square

To determine whether a given vertex ordering σ is a BFS-ordering we first use Algorithm 3 to ensure that σ is a GEN-ordering. We then use Algorithm 4 to determine whether or

³One is included in the other and the two left extremities are different, as are the two right extremities.

not Condition 3 of Theorem 6.2 is satisfied and thus whether or not σ is a BFS-ordering. As with Algorithm 4, we assume that $ln(\sigma(i))$ has been computed for all i between 1 and n .

Algorithm 4: BFS-check

```

 $min \leftarrow n;$     % { $min$  will store the index of the current leftmost value of
                   $ln(\sigma(j))$  for all  $i \leq j \leq n$ .} %
for  $i \leftarrow n$  downto 1 do
    if  $ln(\sigma(i)) > min$  then
         $\perp$  return "NO"
    if  $ln(\sigma(i)) \neq -1$  then
         $\perp$   $min \leftarrow ln(\sigma(i));$ 
return "YES"

```

Theorem 6.3. *Given a GEN-ordering σ , the BFS-check algorithm correctly determines whether σ is a BFS-ordering in $O(n)$ time. The recognition of a BFS-ordering can be done in $O(n + m)$ time.*

Proof. If the algorithm reports that σ is not a BFS-ordering, then consider the triple of vertices $\sigma(min), \sigma(i), \sigma(k)$, where k is the value of i when min was determined. Note that $\sigma(i)$ is not adjacent to $\sigma(min)$ or to any vertices to the left of $\sigma(min)$ and thus this triple forms a forbidden triple as stipulated in Condition 2 of Theorem 4.4.

Now assume that the algorithm reports that σ is a BFS-ordering but for sake of contradiction there exists a forbidden triple on vertices $a <_{\sigma} b <_{\sigma} c$. We let $a' = \sigma(ln(c))$ and note that since b has no neighbours to the left of or equal to a , b is not adjacent to a' or to any vertices to its left.

Thus when $i = \sigma^{-1}(c)$ the algorithm would have reported that σ is not a BFS-ordering. The complexity argument is the same as in the proof of Theorem 6.1. \square

Concerning this particular result on BFS, when the graph is connected it provides as a corollary a linear time algorithm to certify a shortest path between the vertices $\sigma(1)$ and $\sigma(n)$. So in the spirit of [16], this can be used for certifying BFS-based diameter algorithms (see [1, 10]).

6.3 DFS

We now consider DFS and define $Lmax(x)$ for every vertex $x \in V$ to be the rightmost left neighbour of x in σ ; if x has no left neighbours then by convention $Lmax(x) = -1$. The interval $RLeft(x)$ is defined to be $[Lmax(x), x]$; again by convention, if $Lmax(x) = -1$ $RLeft(x)$ is reduced to $[x]$.

Theorem 6.4. *Let $G = (V, E)$ be a graph, and let σ be an ordering of V . Vertex ordering σ is a DFS-ordering of G if and only if*

1. σ is a GEN-ordering of G
2. no two intervals $Right(x)$ and $RLeft(y)$, with $x \neq y$, strictly overlap as intervals.

Proof. \Rightarrow First, notice that every DFS-ordering is also a GEN-ordering. Then, assume, for contradiction, that σ is a DFS-ordering of G , but that in σ $Right(x)$ and $RLeft(y)$ overlap for some $x \neq y$. Necessarily $x <_\sigma y$ and $Lmax(y) <_\sigma x <_\sigma y <_\sigma rn(x)$. $Lmax(y) <_\sigma x$ implies $xy \notin E(G)$. But then the triple $(x, y, rn(x))$ violates the 4-points condition of σ , since y has no neighbour between x and y in σ .

\Leftarrow Assume that σ respects both conditions of the theorem but σ is not a DFS-ordering. Consider a triple (a, b, c) of vertices such that: $a <_\sigma b <_\sigma c$ with $ac \in E$ and $ab \notin E$ but there is no neighbour of b between a and b in σ . Since σ is supposed to be a GEN-ordering, $ac \in E$ implies that b has a neighbour d left to it in σ , which by the above argument, must be before a . Thus $Lmax(b) <_\sigma a$ and therefore the intervals $RLeft(b), Right(a)$ strictly overlap, a contradiction. □

Corollary 6.5. *DFS-orderings can be recognized in $O(n + m)$.*

Proof. Verifying that σ is a generic-ordering can be done in $O(n + m)$ time using Theorem 6.1. To check the second condition, it suffices to build the family of $2n$ intervals and apply a simple 2 states stack automaton [14] to check the overlapping in $O(n)$ time. □

6.4 LBFS and LDFS

Theorem 6.6. *LBFS and LDFS-orderings can be recognized in $O(n(n + m))$ time.*

Proof. To build the recognition algorithm we use the third condition of the relevant theorems in Section 4, in particular 4.6 (LBFS) and 4.7 (LDFS). Both of these conditions are pattern-conditions. The certificate is stored in a table whose entries are keyed by the pair (b, c) where $b <_\sigma c$ and the information will either be the vertex a , where $a <_\sigma b$ that satisfies the corresponding condition or an error message indicating that the condition has been violated. For LBFS and LDFS, the pattern-condition examines a , the leftmost (LBFS) or the rightmost (LDFS) vertex of $N(b) \triangle N(c)$ and requires that $a \in N(b) - N(c)$. It is easy to show that this can be accomplished in time $O(|N(b)| + |N(c)|)$, for any b and c . In all cases, if a satisfies the membership condition then it is stored in the (b, c) 'th entry of the table; otherwise "error" is stored.

Regarding complexity considerations, the table uses $O(n^2)$ space complexity. For the lexicographic searches, the timing requirement is bounded by $\sum_{b \in V} \sum_{c \in V} (|N(b)| + |N(c)|)$

to build the table and $O(n^2)$ time to search for an “error” entry, giving an $O(n(n + m))$ time complexity. \square

These results for LBFS and LDFS do not seem to be optimal, but at least they yield a certificate in case of failure. To improve these algorithms we need to find some new characterizations of LBFS- and LDFS-orderings.

7 Implementation issues

We now consider how to compute an TBLS search, in the case where \prec is a total order. In such a case, at each step of the search, the labels yield a total preorder on the vertices. Such a total preorder (also called weak-order using ordered sets terminology) can be efficiently represented using ordered partitions as can be seen in the next result.

Theorem 7.1. *TBLS(G, \prec, τ) where \prec is a total order can be implemented to run in $O(n + mT(n) \log n)$ time where the \prec comparison time between two labels is bounded by $O(T(n))$.*

Proof. We use the framework of partition refinement [13]. First we sort the adjacency lists with respect to τ , and consider the following algorithm. The input to the algorithm is a graph $G = (V, E)$, a total order \prec on $P_f(\mathbb{N}^+)$, and an ordering τ of V and the output is the TBLS(G, \prec, τ)-ordering σ of V .

Algorithm 5: Computing a TBLS ordering

```

Let  $\mathcal{P}$  be the partition  $\{V\}$ , where the only part (i.e.,  $V$ ) is ordered with respect to  $\tau$ ;
for  $i \leftarrow 1$  to  $n$  do
    Let Eligible be the part of  $\mathcal{P}$  with the largest label with respect to  $\prec$  and;
    let  $x$  be its first vertex;
    replace Eligible by Eligible  $- \{x\}$  in  $\mathcal{P}$  ;
     $\sigma(i) \leftarrow x$ ;
    Refine( $\mathcal{P}, N(x)$ );

```

The algorithm maintains an (unordered) partition \mathcal{P} of the unnumbered vertices. Each part of \mathcal{P} is an ordered list of vertices. The following two invariants hold throughout the execution of the algorithm:

1. The vertices of each part have the same unique (with respect to parts) label;
2. Inside a part, the vertices are sorted with respect to τ .

The action of Refine(\mathcal{P}, A) is to replace each part $P \in \mathcal{P}$ with two new parts: $P \cap A$ and $P - A$ (ignoring empty new parts). It is possible to maintain the two invariants using the data structure from [13], provided the adjacency lists of G are sorted with respect to

τ . After each refinement, each part of \mathcal{P} therefore contains vertices that are twins with respect to the visited vertices (Invariant 1). Thanks to the second invariant, the chosen vertex is always the first vertex (with respect to τ) of part *Eligible*; i.e., $\sigma(i)$ is indeed x .

For the time complexity, $\text{Refine}(\mathcal{P}, N(x))$ takes $O(|N(x)|)$ time [13], so all refinements take $O(n + m)$ time. The only non-linear step is identifying part *Eligible* among all parts of the partition. Each part has a label (the one shared with all its vertices) used as a key in a Max-Heap. $\text{Refine}(\mathcal{P}, N(x))$ creates at most $|N(x)|$ new parts so there are at most m insertions into the heap. The label of a part does not change over time (but empty parts must be removed). There are no more removal operations than insertion operations, each consisting of at most $\log n$ label comparisons (since there are at most n parts at any time). So we get the $O(n + mT(n) \log n)$ time bound. \square

This complexity is not optimal, since it is well-known and already used in some applications (see for example [9]) that classical searches such as BFS, DFS, LBFS can be implemented within the TBLS framework, i.e., solving the tie-break with a given total ordering τ of the vertices, within the same complexity as their current best implementations. To avoid the $T(n)$ costs and the $\log n$ factor, the trick is simply to use an implementation of the search that uses partition refinement (such an implementation exists for BFS, DFS, and LBFS). If we start with a set ordered via τ , there exists a partition refinement implementation that preserves this ordering on each part of the partition, and the tie-break rule means simply choose the first element of the Eligible part. For LDFS, that best known complexity can also be achieved this way. But for Gen-search, MCS and MNS we do not know yet how to achieve linear time, within the TBLS framework.

8 Concluding remarks

We have focused our study on a new formalism that captures many usual graph searches as well as the commonly used multi-sweep application of these searches. The TBLS formalism allows us to define a generic TBLS-orderings recognition algorithm, and gives us a new point of view of the hierarchy amongst graph searches. The new pattern-conditions for Generic Search, BFS and DFS give us a better way (compared to the pattern-conditions presented in [8]) of certifying whether a given ordering could have been produced by such a search. Furthermore, for LBFS and LDFS we do not have to trust the implementation of the search (which can be complicated) but have presented a simple program that just visits the neighbourhood of the vertices of the graph and stores a small amount of information (see Theorem 6.6). The size of this extra information, however, can be bigger than the size of the input, and it may take longer to compute than the actual time needed to perform the search itself.

The landscape of graph search is quite complex. Graph searches can be clustered using the data structures involved in their best implementations (queue, stack, partition refinement ...). In this paper we have tried a more formal way to classify graph searches. This attempt yields an algebraic framework that could be of some interest.

Clearly being an extension (see section 3) is a transitive relation. In fact \ll structures the TBLS graph searches as \wedge -semilattice. The 0 search in this semi-lattice, denoted by the null search or S_{null} , corresponds to the empty ordering relation (no comparable pairs). At every step of S_{null} the Eligible set contains all unnumbered vertices. Therefore for every τ , $TBLS(G, \prec_{S_{null}}, \tau) = \tau$ and so any total ordering of the vertices can be produced by S_{null} .

The infimum between two searches S, S' can be defined as follows:

For every pair of label sets A, B , we define: $A \prec_{S \wedge S'} B$ if and only if $A \prec_S B$ and $A \prec_{S'} B$.

Clearly every extension of S and S' is an extension of $S \wedge S'$. Similarly S and S' are extensions of $S \wedge S'$.

While being as general as GLS, we feel that TBLS is closer to the pattern-conditions presented in [8], since many of the \prec conditions presented in this paper are a rewriting of their pattern-conditions. Still, there are many variants of the searches we studied that do not fall under the TBLS model, such as layered search. We wonder if a more general search model can be found, that would not only include some of these other common searches but would also retain the simplicity of TBLS.

Acknowledgements:

The authors thank Dominique Fortin for his careful reading and numerous suggestions to improve the paper. The first author wishes to thank the Natural Sciences and Engineering Research Council of Canada for their financial support.

References

- [1] Paolo Boldi and Sebastiano Vigna. Four degrees of separation, really. In *ASONAM*, pages 1222–1227, 2012.
- [2] Andreas Brandstädt, Feodor F. Dragan, and Falk Nicolai. Lexbfs-orderings and powers of chordal graphs. *Discrete Mathematics*, 171(1-3):27–42, 1997.
- [3] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. The MIT Press and McGraw-Hill Book Company, 1989.
- [4] Derek G. Corneil. Lexicographic breadth first search - a survey. In *WG*, volume 3353 of *Lecture Notes in Computer Science*, pages 1–19. Springer, 2004.

- [5] Derek G. Corneil. A simple 3-sweep LBFS algorithm for the recognition of unit interval graphs. *Discrete Applied Mathematics*, 138(3):371–379, 2004.
- [6] Derek G. Corneil, Barnaby Dalton, and Michel Habib. LDFS-based certifying algorithm for the minimum path cover problem on cocomparability graphs. *SIAM J. Comput.*, 42(3):792–807, 2013.
- [7] Derek G. Corneil, Jérémie Dusart, Michel Habib, and Ekkehard Köhler. On the power of graph searching for cocomparability graphs. *in preparation*, 2014.
- [8] Derek G. Corneil and Richard Krueger. A unified view of graph searching. *SIAM J. Discrete Math.*, 22(4):1259–1276, 2008.
- [9] Derek G. Corneil, Stephan Olariu, and Lorna Stewart. The LBFS structure and recognition of interval graphs. *SIAM J. Discrete Math.*, 23(4):1905–1953, 2009.
- [10] Pierluigi Crescenzi, Roberto Grossi, Michel Habib, Leonardo LANZI, and Andrea Marino. On computing the diameter of real-world undirected graphs. *Theor. Comput. Sci.*, 514:84–95, 2013.
- [11] Jérémie Dusart and Michel Habib. A new LBFS-based algorithm for cocomparability graph recognition. *submitted*, 2014.
- [12] Martin Charles Golumbic. *Algorithmic Graph Theory and Perfect Graphs (Annals of Discrete Mathematics, Vol 57)*. North-Holland Publishing Co., Amsterdam, The Netherlands, The Netherlands, 2004.
- [13] Michel Habib, Ross M. McConnell, Christophe Paul, and Laurent Viennot. LEXBFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing. *Theor. Comput. Sci.*, 234(1-2):59–84, 2000.
- [14] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation - (2. ed.)*. Addison-Wesley series in computer science. Addison-Wesley-Longman, 2001.
- [15] Richard Krueger, Geneviève Simonet, and Anne Berry. A general label search to investigate classical graph search algorithms. *Discrete Applied Mathematics*, 159(2-3):128–142, 2011.
- [16] Ross M. McConnell, Kurt Mehlhorn, Stefan Näher, and Pascal Schweitzer. Certifying algorithms. *Computer Science Review*, 5(2):119–161, 2011.
- [17] Daniel Meister. Recognition and computation of minimal triangulations for at-free claw-free and co-comparability graphs. *Discrete Applied Mathematics*, 146(3):193–218, 2005.

- [18] George B. Mertzios and Derek G. Corneil. A simple polynomial algorithm for the longest path problem on cocomparability graphs. *SIAM J. Discrete Math.*, 26(3):940–963, 2012.
- [19] Donald J. Rose, Robert Endre Tarjan, and George S. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM J. Comput.*, 5(2):266–283, 1976.
- [20] Micha Sharir. A strong connectivity algorithm and its applications to data flow analysis. *Computers and Mathematics with Applications*, 7:67–72, 1981.
- [21] D. R. Shier. Some aspects of perfect elimination orderings in chordal graphs. *Discrete Applied Mathematics*, 7:325–331, 1984.
- [22] Robert Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2):146–160, 1972.
- [23] Robert Endre Tarjan and Mihalis Yannakakis. Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs. *SIAM J. Comput.*, 13(3):566–579, 1984.
- [24] Marc Tedder. *Applications of Lexicographic Breadth First Search to modular decomposition, split decomposition, and circle graphs*. PhD thesis, University of Toronto, Toronto, Ontario, Canada, 2011.